

# Phase transitions in the condition number distribution of Gaussian random matrices

Isaac Pérez Castillo<sup>1</sup>, Eytan Katzav<sup>2</sup> and Pierpaolo Vivo<sup>3</sup>

1. *Departamento de Sistemas Complejos, Instituto de Física, UNAM, P.O. Box 20-364, 01000 México D.F., México*

2. *Department of Mathematics, King's College London, Strand, London WC2R 2LS, United Kingdom*

3. *Laboratoire de Physique Théorique et Modèles Statistiques (UMR 8626 du CNRS), Université Paris-Sud, Bâtiment 100, 91405 Orsay Cedex, France*

(Dated: March 6, 2014)

We study the statistics of the condition number  $\kappa = \lambda_{\max}/\lambda_{\min}$  (the ratio between largest and smallest squared singular values) of  $N \times M$  Gaussian random matrices. Using a Coulomb fluid technique, we derive analytically and for large  $N$  the cumulative  $\mathcal{P}[\kappa < x]$  and tail-cumulative  $\mathcal{P}[\kappa > x]$  distributions of  $\kappa$ . We find that these distributions decay as  $\mathcal{P}[\kappa < x] \approx \exp(-\beta N^2 \Phi_-(x))$  and  $\mathcal{P}[\kappa > x] \approx \exp(-\beta N \Phi_+(x))$ , where  $\beta$  is the Dyson index of the ensemble. The left and right rate functions  $\Phi_{\pm}(x)$  are independent of  $\beta$  and calculated exactly for any choice of the rectangularity parameter  $\alpha = M/N - 1 > 0$ . Interestingly, they show a weak non-analytic behavior at their minimum  $\langle \kappa \rangle$  (corresponding to the average condition number), a direct consequence of a phase transition in the associated Coulomb fluid problem. Matching the behavior of the rate functions around  $\langle \kappa \rangle$ , we determine exactly the scale of typical fluctuations  $\sim \mathcal{O}(N^{-2/3})$  and the tails of the limiting distribution of  $\kappa$ . The analytical results are in excellent agreement with numerical simulations.

PACS numbers: 5.40.-a, 02.10.Yn, 02.50.Sk, 24.60.-k

**Introduction** - A classical task in numerical analysis is to find the solution  $\mathbf{x}$  of a linear system  $A\mathbf{x} = \mathbf{y}$ , where in the simplest setting  $A$  is a square  $N \times N$  matrix and  $\mathbf{y}$  is a given column vector. The solution can be formally written as  $\mathbf{x} = A^{-1}\mathbf{y}$  provided that  $A$  is invertible. A very important issue for numerical stability is: how does a small change in the entries of  $\mathbf{y}$  or of  $A$  propagate on the solution  $\mathbf{x}$ ?

The system of equations above is said to be *well-* (or *ill-*) *conditioned* if a small change in the coefficient matrix  $A$  or in the right hand side  $\mathbf{y}$  results in a small (or large) change in the solution vector  $\mathbf{x}$ . An ill-conditioned system produces a solution that cannot be trusted, as numerical inaccuracies in the inputs are amplified and propagated to the output [1].

A standard indicator of the reliability of numerical solutions is the *condition number* (CN)  $\kappa = \lambda_{\max}/\lambda_{\min} \geq 1$ , where  $\lambda_{\min}$  and  $\lambda_{\max}$  are the smallest and largest *squared* singular values of  $A$ , i.e. the positive eigenvalues of  $AA^T$  (the square root  $\sqrt{\kappa}$  of CN is alternatively used very frequently). The quantity  $\ln_b \kappa$  is essentially a worst-case estimate of how many base- $b$  digits are lost in solving numerically that linear system, which is *singular* if  $\kappa$  is infinite, *ill-conditioned* if  $\kappa$  is “too large”, and *well-conditioned* if  $\kappa$  is close to its minimum value 1.

Computing  $\kappa$  for a large coefficient matrix  $A$  in a fast and efficient way can be, however, as difficult a task as solving the original system in the first place [2]. To overcome this problem, Goldstine and von Neumann [3, 4] proposed instead to study the generic features of  $\kappa$  associated to a *random* matrix  $A$  with normally distributed elements [5]. What is the *typical* (expected) CN for a system of size  $N$ ? And what is a sensible estimate for the size of its fluctuations?

Modern applications of a random condition number of more general (rectangular) matrices  $N \times M$  include wireless communication systems [6–10], spectrum sensing algorithms [11–13], convergence rate of iterative schemes [14], compressed sensing [15], finance [16], meteorology [17] and performance

assessment of principal component analysis [18] among others.

The statistics of  $\sqrt{\kappa}$  was first computed by Edelman [19] for  $2 \times M$  random Gaussian matrices, as well as the limiting distribution of  $\sqrt{\kappa}/N$  for large  $N \times N$  matrices. The rectangular case was recently considered in [20]. Different bounds for the tails were given in [21–23]. Exact formulae for the distribution of  $\kappa$  for finite  $N, M$  also exist in terms of cumbersome series of zonal polynomials [24, 25] or an integral of a determinant [26], whose evaluation becomes impractical even for moderate matrix sizes. Approximate results for correlated non-central Gaussian matrices can be found e.g. in [27]. Other definitions for the CN also exist [5, 28].

Unfortunately, almost nothing is known about the most dreaded (or welcomed) scenarios for applications, namely the occurrence of *atypical* instances, where the CN is much larger (or smaller) than its expected value. In this Letter, by suitably adapting the Coulomb fluid method of statistical mechanics, we provide an analytical solution to this outstanding problem for large rectangular instances. We show that the large deviation statistics of the CN of Gaussian matrices, expressed in terms of elementary functions, has a rich and elegant structure. As a bonus, we also derive the scale of typical fluctuation of the CN around  $\langle \kappa \rangle$ , and the tails of its limiting distribution. Let us first summarize our setting and main results.

**Summary of results** - We consider rectangular  $N \times M$  ( $M > N$ ) matrices  $A$  with Gaussian distributed entries (real, complex or quaternions, labelled by the Dyson index  $\beta = 1, 2, 4$  respectively, or actually for general  $\beta > 0$  as discussed in [29]). Forming the corresponding  $N \times N$  covariance matrix  $W = AA^T$  [43], which defines the Wishart ensemble [30], we define its rectangularity parameter  $\alpha = M/N - 1 > 0$  and the CN  $\kappa = \lambda_{\max}/\lambda_{\min} > 1$ . Here  $\lambda_{\max}$  and  $\lambda_{\min}$  are the largest and smallest eigenvalues of  $W$ . We consider the cumulative  $\mathcal{P}[\kappa < x]$  and tail-cumulative (also known as exceedance or

survival function)  $\mathcal{P}[\kappa > x]$  distributions of  $\kappa$ , when  $N$  and  $M$  are large and  $\alpha$  is kept finite. Using a Coulomb fluid technique we find that for large  $N$  both distributions obey large deviation laws, namely they decay for large  $N$  as [44]

$$\mathcal{P}[\kappa < x] \approx \exp(-\beta N^2 \Phi_-(x)), \quad (1)$$

$$\mathcal{P}[\kappa > x] \approx \exp(-\beta N \Phi_+(x)). \quad (2)$$

The left and right rate functions  $\Phi_{\pm}(x)$  (depending parametrically on  $\alpha$ , but not on  $\beta$ ) are given in (10) and (12) and plotted in Fig. 1. Both functions are supported on  $x \in (1, \infty)$  and have a minimum (zero) at  $\langle \kappa \rangle = [(1 + \sqrt{1 + \alpha})/(1 - \sqrt{1 + \alpha})]^2 > 1$ . Therefore the corresponding *density* of  $\kappa$  is peaked around  $\langle \kappa \rangle$ , which is precisely its mean value for large  $N$  [45]. Crossing  $\langle \kappa \rangle$ , both functions freeze to the zero value, and around  $\langle \kappa \rangle$  they have an interesting non-analytic behavior, characterized by a third-order (for  $\Phi_-(x)$ ) [31] and second-order (for  $\Phi_+(x)$ ) discontinuity. Both these non-analytic behaviors and the different scaling with  $N$  between (1) and (2) are direct consequences of *freezing phase transitions* [32] in an associated Coulomb fluid problem. The physics of the two branches is however entirely different (see below for details). Matching the behavior of the rate functions around  $\langle \kappa \rangle$ , we also determine exactly the size ( $\sim \mathcal{O}(N^{-2/3})$ ) of typical fluctuations of  $\kappa$  and the tails of its limiting distribution. We now begin by recalling some well-known facts about Wishart matrices.

**Generalities** - The joint probability density (jpd) of the  $N$  (real and positive) eigenvalues is given by [33, 34]

$$P_{\beta}(\lambda) = \frac{1}{Z_0} e^{-\frac{1}{2} \sum_{i=1}^N \lambda_i} \prod_{i=1}^N \lambda_i^{\frac{\beta}{2}(\alpha N + 1) - 1} \prod_{i < j} |\lambda_i - \lambda_j|^{\beta}, \quad (3)$$

where  $Z_0$  is a normalization constant and  $\beta = 1, 2, 4$  is the Dyson index of the ensemble. Balancing the first and third terms in (3), it is quite easy to estimate that the typical scale of an eigenvalue is  $\sim \mathcal{O}(N)$ . Thus, after rescaling  $\lambda_i \rightarrow \beta N \lambda_i$ , the jpd (3) can be rewritten in the form  $P_{\beta}(\lambda) \propto \exp(-\beta N^2 E[\{\lambda\}])$ , where the  $\mathcal{O}(1)$  energy is

$$E[\{\lambda\}] = \frac{1}{2N} \sum_{j=1}^N \lambda_j - \frac{\alpha}{2N} \sum_{j=1}^N \ln \lambda_j - \frac{1}{2N^2} \sum_{j \neq k} \ln |\lambda_j - \lambda_k|. \quad (4)$$

Written in this form, the jpd (3) resembles the Gibbs-Boltzmann canonical weight of a 2D fluid of charged particles, confined on the semi-infinite (positive) line and in equilibrium at inverse temperature  $\beta$  under competing interactions: the external linear-logarithmic potential in (4) drives the charges towards its minimum, while the third term (representing an all-to-all repulsive interaction of the Coulomb type in 2D) spreads them apart. This thermodynamical analogy, originally pioneered by Dyson [35], has been lately employed in several different contexts [31, 36–40].

The average spectral density of the Wishart model  $\rho(\lambda) = N^{-1} \sum_{i=1}^N \langle \delta(\lambda - \lambda_i) \rangle$  (where  $\langle \dots \rangle$  denotes averaging with

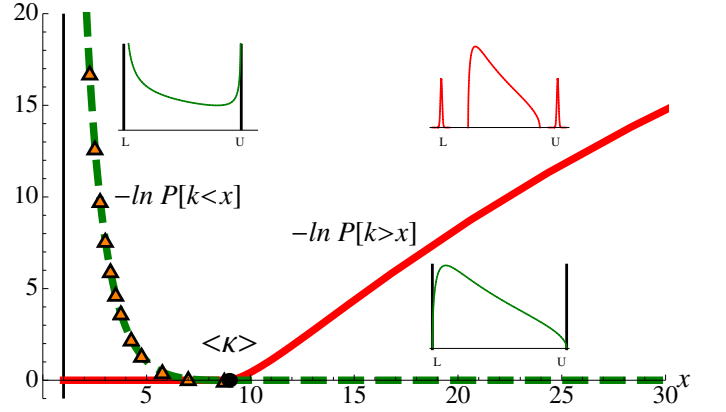


FIG. 1: Plot of  $-\ln \mathcal{P}[\kappa < x]$  (dashed green, Eq. (1)) and  $-\ln \mathcal{P}[\kappa > x]$  (solid red, Eq. (2)), together with numerical simulations for the left branch [42]. The two rate functions  $\beta N^2 \Phi_-(x)$  and  $\beta N \Phi_+(x)$  freeze to the zero value upon crossing  $\langle \kappa \rangle$ . The insets describe the corresponding phases of the Coulomb fluid (active vs. inactive barriers for  $\Phi_-(x)$  and the pulling of individual extreme charges for  $\Phi_+(x)$ ). An arbitrary value of  $N$  has been chosen to produce a reasonably looking plot in which the two branches are visible.

respect to the jpd (3)) is expected for large  $N$  to have the scaling form  $\rho(\lambda) = N^{-1} \rho_{\text{mp}}(\lambda/N)$ , where the function  $\rho_{\text{mp}}(x) = \frac{1}{2\pi x} \sqrt{(x - z_-)(z_+ - x)}$  is the celebrated Marčenko-Pastur (MP) law on the compact support (for  $\alpha > 0$ )  $x \in [z_-, z_+]$  with  $z_{\pm} = (1 \pm \sqrt{1 + \alpha})^2$ . This MP law is a particular case of the general solution (9) of the integral equation (8) below (see Fig. 2, top) when the two barriers  $L, U$  are ineffective ( $L \leq z_-$  and  $U \geq z_+$ ). We start now by considering the cumulative distribution of the CN first, and get to the tail-cumulative afterwards.

**Cumulative distribution** - The cumulative distribution  $\mathcal{P}[\kappa < x]$  of the CN  $\kappa$  (depending parametrically on  $\beta$  and  $\alpha = M/N - 1 > 0$ ), can be written as [26, 42]

$$\mathcal{P}[\kappa < x] = \frac{1}{(N-1)!} \int_0^{\infty} d\lambda_1 \left[ \int \dots \int_{\lambda_1}^{x\lambda_1} \prod_{j=2}^N d\lambda_j P_{\beta}(\lambda) \right]. \quad (5)$$

The goal is to evaluate this multiple integral for large  $N$  by the Coulomb fluid method. The first step is to rewrite the jpd (3) in the Gibbs-Boltzmann form described above. Here,  $N - 1$  fluid particles are however not free to spread on the whole positive line, but instead *constrained* to live within the box  $[\lambda_1, x\lambda_1]$ , where  $\lambda_1$  is the (free) position of the leftmost particle.

The second step consists of a coarse-graining procedure, where one introduces a normalized density of particles  $\rho(\lambda) = (N - 1)^{-1} \sum_{i=2}^N \delta(\lambda - \lambda_i)$  for the  $N - 1$  particles  $\lambda_i$  ( $i \neq 1$ ) living *inside* the box. Using the replacement rule  $\sum_{i>1} g(\lambda_i) = (N - 1) \int d\lambda \rho(\lambda) g(\lambda)$ , we can convert the energy function  $E[\{\lambda\}]$  into a continuous action  $S$  (depending on  $\rho$ , and parametrically on the location of the leftmost particle  $\lambda_1$  and  $x$ ). The multiple integration (5) is therefore

interpreted as the canonical partition function of the associated Coulomb fluid, where the sum over all microscopic configurations of  $\{\lambda\}$  compatible with the normalized density  $\rho$  amounts to a functional integration over  $\rho$  and a standard integration over  $\lambda_1$ . Eventually, these resulting integrals are evaluated using the saddle point method. Performing these steps, we get

$$\mathcal{P}[\kappa < x] \propto \int_0^\infty d\xi \int \mathcal{D}[\rho, C] e^{-\beta N^2 S[\rho, C; \xi, x\xi]}, \quad (6)$$

where we renamed  $\lambda_1 \rightarrow \xi$  for later convenience and the action of this fluid (confined between the lower  $L$  and upper  $U$  barriers of the box) is

$$S[\rho, C; L, U] = \int_L^U d\lambda \rho(\lambda) V(\lambda) + C - \frac{1}{2} \iint_L^U d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \ln |\lambda - \lambda'|. \quad (7)$$

Here  $V(\lambda) = (\lambda - \alpha \ln \lambda)/2 - C$  and  $C$  is a Lagrange multiplier enforcing normalization of  $\rho$ . Eq. (7) is easily identified as the continuous version of the energy Eq. (4), where we have neglected subleading  $\mathcal{O}(N)$  contributions [42].

Evaluating the functional integral in (6) by the saddle point method,  $\left. \frac{\delta S}{\delta \rho} \right|_{\rho=\rho^*} = 0$ , we get

$$V(\lambda) = \int_L^U d\lambda' \rho^*(\lambda') \ln |\lambda - \lambda'|, \quad (8)$$

where the solution  $\rho^*(\lambda)$  is just the equilibrium density of the Coulomb fluid constrained to live within the box  $[L, U]$ . Clearly, if we release the barriers  $L, U$  we expect to recover the unconstrained MP law,  $\rho^*(\lambda) \rightarrow \rho_{\text{mp}}(\lambda)$ . Solving this integral equation for a normalized  $\rho^*$  between two barriers at  $L$  and  $U$  is one of the main technical challenges that we managed to overcome. Skipping details [42], we find that the general solution of (8) is [41]

$$\rho^*(\lambda) = \frac{(x_+(L, U) - \lambda)(\lambda - x_-(L, U))}{2\pi\lambda\sqrt{(U - \lambda)(\lambda - L)}} \mathbb{1}_{[L, U]}(\lambda), \quad (9)$$

where  $x_+(L, U) \geq U > L \geq x_-(L, U)$ ,  $x_\pm(L, U)$  are the roots of  $x^2 - x\left(\frac{L+U}{2} + \alpha + 2\right) + \alpha\sqrt{LU} = 0$ , and  $\mathbb{1}_{[a, b]}(x) = 1$  if  $x \in [a, b]$  and 0 otherwise.

How does this density look like for given values of  $L$  and  $U$ ? Four different shapes (phases of the fluid) are possible [42] for  $\alpha > 0$  that are plotted in Fig. 2 (top). For example, setting  $(L, U) = (z_-, z_+)$  the corresponding density (9) is the MP law  $\rho^*(x) = \rho_{\text{mp}}(x)$  (phase II). This critical MP point, which is marked in Fig. 2 (bottom), separates region II, where the barriers are *ineffective* ( $L < z_-$  and  $U > z_+$ ) and the equilibrium density is again just  $\rho_{\text{mp}}(x)$ , from region I (where the barriers are instead *effective* in compressing the MP sea).

Once we have evaluated the functional integral by the saddle-point method (which implies inserting the density (9) in

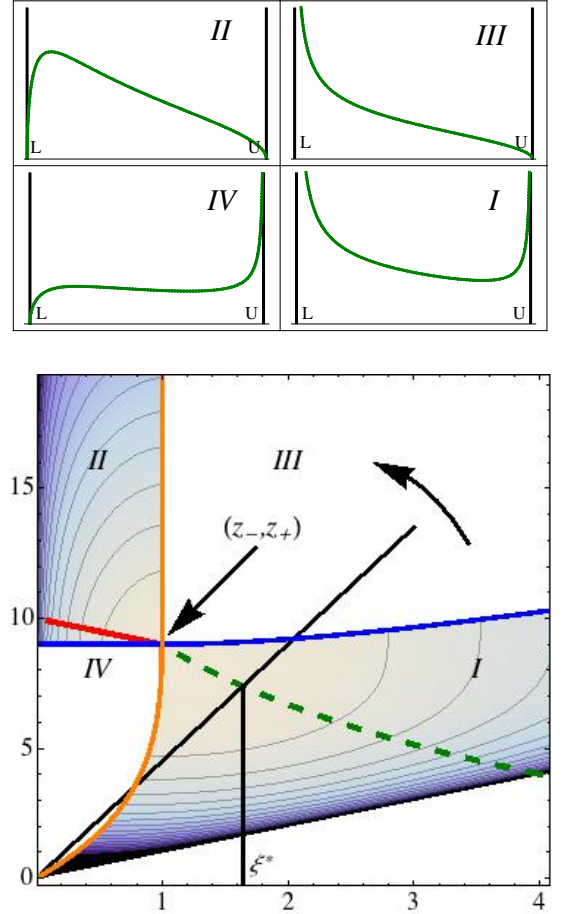


FIG. 2: Top: The four phases of the fluid. Region I: the two barriers compress effectively the fluid. Region II: MP law, where the barriers do not affect the fluid. Region III and IV: only the lower or the upper barrier is active, respectively (this scenario is not realized in our CN setting). The analytical expressions in regions III and IV were first derived in [38, 40], respectively.

Down: Regions in the  $(L, U)$  plane where the density Eq. (9) has different shapes, according to top panels, for  $\alpha = 3$ . We plot level curves of the action (7)  $S[\rho^*, C; L, U]$  (I) and energy difference  $\Delta e(L, U)$  [Eq. (0.15) in [42]] (IV). On the dashed green and solid red extremal lines, the action and the energy difference are minimal, respectively. The abscissas  $\xi^*$  (solution of the saddle point equations in the two cases) are given by the intersection of the straight line  $U = xL$  (solid black, with the left arrow pointing in the direction of increasing slope  $x$ ) with such extremal lines. Besides, the solid orange line corresponds to the condition  $x_-(L, U) = L$  along which the lower barrier is ineffective (and the upper is effective), while the solid blue line corresponds to the condition  $x_+(L, U) = U$  along which the upper barrier is ineffective (and the lower is effective).

the action (7)) we set  $L = \xi$  and  $U = x\xi$  and evaluate the remaining  $\xi$ -integral again by the saddle-point method. This yields an optimal value  $\xi^*$  as the solution of  $\left. \frac{d}{d\xi} S[\rho^*, C; \xi, x\xi] \right|_{\xi=\xi^*} = 0$ . This value  $\xi^*(x)$  is marked in Fig. 2 (bottom) as the intersection between the straight line  $U = xL$  of varying slope  $x > 1$  and the dashed green line on which the action  $S[\rho^*, C; L, U]$  is minimal.

The final result reads  $\mathcal{P}[\kappa < x] \approx e^{-\beta N^2 \Phi_-(x)}$ , where the  $\mathcal{O}(N^2)$  decay is traced back to the high energy cost in compressing the whole sea of strongly correlated particles. Here  $\Phi_-(x) = S[\rho^*, C; \xi^*, x\xi^*] - S[\rho_{\text{mp}}, C; z_-, z_+]$ , where the second term comes from the normalization factor and needs to be subtracted. We eventually obtain

$$\Phi_-(x) = \frac{1}{8} \left[ f_1^{(\alpha)}(1 + \sqrt{x}) + \ln f_2^{(\alpha)}(1 + \sqrt{x}) \right] \mathbb{1}_{(1, \langle \kappa \rangle)}(x), \quad (10)$$

where  $f_{1,2}^{(\alpha)}(\omega)$  are elementary functions listed in [42]. The rate function  $\Phi_-(x)$  thus freezes to the value 0 as  $x$  increases up to the critical value  $\langle \kappa \rangle$  (implying  $\xi^*(\langle \kappa \rangle) \rightarrow z_-$ ). Beyond this limit, the barriers are no longer effective, and new physical insights are needed to tackle the tail-cumulative regime (see next subsection). The limits are  $\Phi_-(x \rightarrow \langle \kappa \rangle^-) \sim K(\alpha)(\langle \kappa \rangle - x)^3$  and  $\Phi_-(x \rightarrow 1^+) \sim (-1/2) \ln(x-1)$ , where  $K(\alpha) = -(-1 + \sqrt{1+\alpha})^8 / 96\sqrt{1+\alpha}(1 + \sqrt{1+\alpha})^4$ . This implies a third order discontinuity across  $\langle \kappa \rangle$  as anticipated. Also, close to 1, the density of  $\kappa$  has the following power-law tail,  $\mathcal{P}[\kappa = 1 + \epsilon] \sim \epsilon^{\beta N^2/2}$  to leading order in  $N$  for  $\epsilon \rightarrow 0^+$ . Although formally valid only for  $\alpha > 0$ , it turns out that in the limit  $\alpha \rightarrow 0$  (square Gaussian matrices, where the scaling with  $N$  is different) the rate function (10) is still well-behaved and we recover Edelman's result [19] to leading order in  $N$  [42]. We now turn to the tail-cumulative distribution (the right branch in Fig. 1).

*Tail-cumulative distribution* - Contrary to the previous case, the tail-cumulative distribution  $\mathcal{P}[\kappa > x]$  does *not* admit a multiple-integral representation of the type (5), which could be mapped to the physics of a fluid “trapped” between two hard barriers.

The starting point of the calculation is again the energy function (4), though. The Coulomb fluid physics suggests that atypically large values of the CN  $\kappa = \lambda_{\text{max}}/\lambda_{\text{min}}$  are obtained when the right-and-leftmost particles are *pulled* away from the MP sea in opposite directions ( $\lambda_{\text{max}} - \lambda_{\text{min}} \sim \mathcal{O}(N)$ ), a procedure that is energetically not able to generate macroscopic rearrangements *within* the MP sea. This elegant energetic argument was first introduced in [39]. Following this physical picture, the right rate function  $\Phi_+(x)$  is determined by the  $\mathcal{O}(N)$  energy cost  $\Delta E(L, U)$  in pinning the leftmost and rightmost charges at  $L$  and  $U$ , well outside the *unperturbed* MP sea in between (see red inset in Fig. 1). The level curves of the  $\mathcal{O}(1)$   $\Delta e(L, U) := \Delta E(L, U)/\beta N$  are depicted in region II of Fig. 2, together with the extremal line (solid red) where it attains its minimum value. Setting now  $L = \xi$  and  $U = \xi x$ , the energetically most favored position  $\xi^*$  for the leftmost outlier will be determined again by the intersection point of that extremal line and the straight line  $U = xL$ .

Skipping details [42], this change in energy can be written for

large  $N$  as

$$\Delta e(\xi, \xi x) = \frac{(\xi - z_-) + (\xi x - z_+)}{2} - \frac{\alpha}{2} \ln \frac{\xi^2 x}{z_- z_+} - \int_{z_-}^{z_+} d\eta \rho_{\text{mp}}(\eta) \ln \left| \frac{(\xi - \eta)(\eta - \xi x)}{(z_- - \eta)(\eta - z_+)} \right|. \quad (11)$$

The probability of this “pinned” configuration of eigenvalues (yielding a CN  $\kappa$  *exactly* equal to  $x$ ) is  $\propto \exp(-\beta N \Delta e(\xi, \xi x))$ . Finding the optimal position  $\xi^*$  for the leftmost particle by minimizing (11) with respect to  $\xi$ , we eventually obtain  $\mathcal{P}[\kappa > x] \approx \exp(-\beta N \Phi_+(x))$ , where  $\Phi_+(x) = \Delta e(\xi^*(x), \xi^*(x)x)$  is given by [46]

$$\Phi_+(x) = \ln \left[ \left( g^{(\alpha)}(x) \right)^{\alpha/2} \left( h^{(\alpha)}(x) \right)^{2(2+\alpha)} \right] \mathbb{1}_{(\langle \kappa \rangle, \infty)}(x). \quad (12)$$

The functions  $g^{(\alpha)}(x)$  and  $h^{(\alpha)}(x)$  have lengthy but explicit expressions in terms of elementary functions [42]. The rate function  $\Phi_+(x)$  again freezes to the value 0 as  $x$  decreases down to the critical value  $\langle \kappa \rangle$  (implying  $\xi^*(\langle \kappa \rangle) \rightarrow z_-$ ), where the pinned outliers reconnect with the MP sea. The limits are  $\Phi_+(x \rightarrow \langle \kappa \rangle^+) \sim J(\alpha)(x - \langle \kappa \rangle)^{3/2}$  and  $\Phi_+(x \rightarrow \infty) \sim (\alpha/2) \ln x$ , where  $J(\alpha) = \sqrt{2} \sqrt[4]{\alpha+1} (\sqrt{\alpha+1} - 1)^4 / 3\sqrt{\alpha+2} (\sqrt{\alpha+1} + 1)^2$ . This implies a second order discontinuity across  $\langle \kappa \rangle$  as anticipated. Also, at infinity the density of  $\kappa$  therefore decays as a power-law,  $\mathcal{P}[\kappa = x] \sim x^{-\alpha\beta N/2}$  to leading order in  $N$  for  $x \rightarrow \infty$ .

*Conclusions* - In summary, we have computed analytically for large  $N$  the cumulative and tail-cumulative distributions of the CN  $\kappa$  of rectangular  $N \times M$  Gaussian random matrices. Mapping the problem to the calculation of the free energy of an associated Coulomb fluid, we found that the random variable  $\kappa$  satisfies large deviation laws governed by rate functions  $\Phi_{\pm}(x)$  to the left and to the right of the expected value  $\langle \kappa \rangle$ , albeit with different scalings with  $N$ . These rate functions (see (10) and (12)) essentially describe the probability of sampling a random Gaussian matrix with an atypically large (or small) CN. They are monotonic and convex, with a non-analytic behavior at their zero  $\langle \kappa \rangle$ . This non-analytic behavior is a direct consequence of a very rich thermodynamics of the associated Coulomb fluid. More precisely, across  $\langle \kappa \rangle$  a *compressed* (left branch) or *stretched* (right branch) fluid undergoes *freezing* phase transitions of different orders (third and second respectively).

Matching the behavior of the rate functions close to their minimum  $\langle \kappa \rangle$ , we deduce that *typical* fluctuations of  $\kappa$  around  $\langle \kappa \rangle$  should occur on a scale of  $\mathcal{O}(N^{-2/3})$ , and setting  $\chi = f(\alpha)N^{2/3}(\kappa - \langle \kappa \rangle)$ , with  $f(\alpha) = 2^{1/3}(1 + \alpha)^{1/6}(-1 + \sqrt{1 + \alpha})^{8/3} / (1 + \sqrt{1 + \alpha})^{4/3}$ , the scaled random variable  $\chi$  has a  $N$ - and  $\alpha$ -independent distribution  $\mathcal{P}[\chi < x] = \mathcal{F}_{\beta}(x)$  with tails  $\sim \exp(-\beta|x|^3)$  for  $x \rightarrow -\infty$  and  $\sim \exp(-\beta x^{3/2})$  for  $x \rightarrow \infty$ . The scaling is in agreement with a recent result [20], valid only for  $M \gg N^3$ . Our analytical results have been numerically checked with excellent agreement [42].

*Acknowledgments* - PV acknowledges financial support from Labex-PALM (project RandMat). We thank G. Schehr and F. D. Cunden for helpful discussions. IPC acknowledges hospitality by the LPTMS (Univ. Paris Sud) where this work was completed.

- 
- [1] S. Smale, *On the efficiency of algorithms of analysis*, Bull. (N. S.) Amer. Math. Soc. **13**, 87 (1985).
  - [2] H. Avron, A. Druinsky, and S. Toledo, *Iterative spectral condition-number estimation*, [arXiv: 1301.1107] (2013).
  - [3] H. H. Goldstine and J. von Neumann, *Numerical inverting of matrices of high order*, Amer. Math. Soc. Bull. **53**, 1021 (1947).
  - [4] H. H. Goldstine and J. von Neumann, *Numerical inverting of matrices of high order, II*, Amer. Math. Soc. Proc. **2**, 188 (1951).
  - [5] J. W. Demmel, *The probability that a numerical analysis problem is difficult*, Mathematics of Computation **50**, 449 (1988).
  - [6] V. Erceg, P. Soma, D. S. Baum, and A. J. Paulraj, *Capacity obtained from multiple-input multiple-output channel measurements in fixed wireless environments at 2.5 GHz*, in Proc. IEEE Int. Conf. Commun. (ICC), vol. 1, New York, pp. 396-400 (2002).
  - [7] M. D. Batarieri et al., *Wideband MIMO mobile impulse response measurements at 3.7 GHz*, in Proc. IEEE Veh. Technol. Conf. (VTC), Birmingham, AL, pp. 26-30 (2002).
  - [8] H. Artes, D. Seethaler, and F. Hlawatsch, *Efficient detection algorithms for MIMO channels: a geometrical approach to approximate ML detection*, IEEE Trans. Signal Process. **51**, 2808 (2003).
  - [9] D. Wubben, R. Bohnke, V. Kuhn, and K. D. Kammeyer, *MMSE-based lattice-reduction for near-ML detection of MIMO systems*, in Proc. ITG Work. Smart Antennas (WSA), Munich, Germany, pp. 106-113 (2004).
  - [10] J. Maurer, G. Matz, and D. Seethaler, *Low-complexity and full-diversity MIMO detection based on condition number thresholding*, in Proc. IEEE Int. Conf. Acoustics Speech Signal Process. (ICASSP), vol. 3, Honolulu, HI, pp. 61-64 (2007).
  - [11] Y. Zeng and Y.-C. Liang, *Eigenvalue based spectrum sensing algorithms for cognitive radio*, IEEE Trans. Commun. **57**, 1784 (2009).
  - [12] F. Penna, R. Garello, and M. Spirito, *Cooperative spectrum sensing based on the limiting eigenvalue ratio distribution in Wishart matrices*, IEEE Commun. Lett. **13**, 507 (2009).
  - [13] L. S. Cardoso, M. Debbah, P. Bianchi, and J. Najim, *Cooperative spectrum sensing using random matrix theory*, in Proc. IEEE Int. Symp. Wireless Pervasive Comput. (ISWPC), Santorini, Greece, pp. 334-338 (2008).
  - [14] G. Strang, *Linear Algebra and its Applications*, 3rd edition. San Diego: Harcourt Brace Jovanovich Inc. (1988).
  - [15] H. Rauhut, *Compressive Sensing and Structured Random Matrices*, in *Theoretical Foundations and Numerical Methods for Sparse Recovery*, Berlin, Boston: DE GRUYTER, 1-92 (2010).
  - [16] A. Binder and M. Aichinger, *A Workout in Computational Finance*, The Wiley Finance Series, John Wiley & Sons (2013).
  - [17] M. Piringer, M. Hrad, S. Stenzel, and M. Huber-Humer, *Reconstruction of CH<sub>4</sub> emissions from a biogas plant - meteorological aspects*, Proceedings of 15th conference on "Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes", Madrid, Spain, May 6-9, 2013 [online at <http://www.harmo.org/conferences/Madrid/15harmo.asp>].
  - [18] A. Kelemen, A. Abraham, and Y. Liang, *Computational Intelligence in Medical Informatics*, Vol. 85 of Studies in Computational Intelligence, Springer (2008).
  - [19] A. Edelman, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl. **9**, 543 (1988).
  - [20] T. Jiang and D. Li, *Approximation of rectangular Beta-Laguerre ensembles and large deviations*, J. Theor. Prob. **1-44** (2013).
  - [21] A. Edelman and B. D. Sutton, *Tails of condition number distributions*, SIAM J. Matrix Anal. Appl. **27**, 547 (2005).
  - [22] Z. Chen and J. J. Dongarra, *Condition numbers of Gaussian random matrices*, SIAM J. Matrix Anal. Appl. **27**, 603 (2005).
  - [23] J.-M. Azaïs and M. Wschebor, *Upper and lower bounds for the tails of the distribution of the condition number of a Gaussian matrix*, SIAM J. Matrix Anal. Appl. **26**, 426 (2005).
  - [24] W. Anderson and M. T. Wells, *The exact distribution of the condition number of a Gaussian matrix*, SIAM J. Matrix Anal. Appl. **31**, 1125 (2009).
  - [25] T. Ratnarajah, R. Vaillancourt, and M. Alvo, *Eigenvalues and condition numbers of complex random matrices*, SIAM J. Matrix Anal. Appl. **26**, 441 (2005).
  - [26] M. Matthaiou, M. R. McKay, P. J. Smith, and J. A. Nossek, *On the condition number distribution of complex Wishart matrices*, IEEE Transactions on Communications **58**, 1705 (2010).
  - [27] L. Wei and O. Tirkkonen, *Approximate condition number distribution of complex non-central correlated Wishart matrices*, IEEE International Conference on Communications (ICC), pp.1-5, 5-9 June 2011.
  - [28] A. Edelman, *On the distribution of a scaled condition number*, Math. of Comp. **58**, 185 (1992).
  - [29] I. Dumitriu and A. Edelman, *Matrix models for Beta ensembles*, J. Math. Phys. **43**, 5830 (2002).
  - [30] J. Wishart, *The generalised product moment distribution in samples from a normal multivariate population*, Biometrika **20A**, 32 (1928).
  - [31] S. N. Majumdar and G. Schehr, *Top eigenvalue of a random matrix: large deviations and third order phase transition*, J. Stat. Mech.: Theory and Experiment **P01012** (2014).
  - [32] C. Texier and S. N. Majumdar, *Wigner time-delay distribution in chaotic cavities and freezing transition*, Phys. Rev. Lett. **110**, 250602 (2013).
  - [33] R. A. Fisher, *The sampling distribution of some statistics obtained from non-linear equations*, Annals of Eugenics **9**, 238 (1939).
  - [34] P. L. Hsu, *On the distribution of the roots of certain determinantal equations*, Annals of Eugenics **9**, 250 (1939).
  - [35] F. J. Dyson, *Statistical Theory of the Energy Levels of Complex Systems. I*, J. Math. Phys. **3**, 140 (1962); *Statistical Theory of the Energy Levels of Complex Systems. II*, J. Math. Phys. **3**, 157 (1962); *Statistical Theory of the Energy Levels of Complex Systems. III*, J. Math. Phys. **3**, 166 (1962).
  - [36] D. S. Dean and S. N. Majumdar, *Extreme value statistics of eigenvalues of Gaussian random matrices*, Phys. Rev. E **77**, 041108 (2008).
  - [37] D. S. Dean and S. N. Majumdar, *Large deviations of extreme eigenvalues of random matrices*, Phys. Rev. Lett. **97**, 160201 (2006).
  - [38] P. Vivo, S. N. Majumdar, and O. Bohigas, *Large deviations of the maximum eigenvalue in Wishart random matrices*, J. Phys. A: Math. Theor. **40**, 4317 (2007).

- [39] S. N. Majumdar and M. Vergassola, *Large deviations of the maximum eigenvalue for Wishart and Gaussian random matrices*, Phys. Rev. Lett. **102**, 060601 (2009).
- [40] E. Katzav and I. Pérez Castillo, *Large deviations of the smallest eigenvalue of the Wishart-Laguerre ensemble*, Phys. Rev. E **82**, 041004 (2010).
- [41] F. G. Tricomi, *Integral Equations* (Pure Appl. Math V, Interscience, London, 1957).
- [42] See Supplementary Material.
- [43] Here  $^T$  stands for transpose ( $\beta = 1$ ), hermitian conjugate ( $\beta = 2$ ) and symplectic conjugate ( $\beta = 4$ ).
- [44] Here  $\approx$  stands for the logarithmic equivalence  $\lim_{N \rightarrow \infty} -\ln \mathcal{P}[\kappa < x]/\beta N^2 = \Phi_-(x)$  and similarly for the tail-cumulative branch.
- [45] The typical value  $\langle \kappa \rangle$  for large  $N$  is just the ratio of the average  $\langle \lambda_{\max} \rangle = (1 + \sqrt{1 + \alpha})^2$  and the average  $\langle \lambda_{\min} \rangle = (1 - \sqrt{1 + \alpha})^2$ .
- [46] Note that this energetic argument gives (strictly speaking) the *density* of  $\kappa$ ,  $\mathcal{P}[\kappa = x]$  and not its tail-cumulative distribution  $\mathcal{P}[\kappa > x]$ . However, the large  $N$  decay of the two is the same to leading order.

## SUPPLEMENTARY MATERIAL

### Contents

<b>Supplementary Material</b>	6
Derivation of the expression for the action in Eq. (7)	6
Solution of integral equation Eq. (8)	7
Left rate function	8
Square case: Edelman's result	9
Right rate function	10
Order of phase transitions	11
Numerical simulations for the left rate function	11

### Derivation of the expression for the action in Eq. (7)

Let us start by noticing that  $\mathcal{P}[\kappa < x] = \text{Prob}[\lambda_{\max} < x\lambda_{\min}]$ , which can be worked out as follows

$$\begin{aligned}
 \mathcal{P}[\kappa < x] &= \text{Prob}(\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N \leq \kappa \lambda_1) \\
 &= \int_0^\infty dx \text{Prob}(x \leq \lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_{N-1} \leq \lambda_N \leq \kappa x, \lambda_1 = x) \\
 &= \frac{1}{(N-1)!} \int_0^\infty dx \text{Prob}(\lambda_1 = x, \{\lambda_1 \leq \lambda_i \leq \kappa \lambda_1, i = 2, \dots, N\})
 \end{aligned}$$

or in terms of the jpd of eigenvalues of the Wishart ensemble we write

$$\mathcal{P}[\kappa < x] = \frac{1}{(N-1)!} \int_0^\infty d\lambda_1 \left[ \int_{\lambda_1}^{x\lambda_1} d\lambda_2 \dots \int_{\lambda_1}^{x\lambda_1} d\lambda_N P_\beta(\boldsymbol{\lambda}) \right]$$

To go to a continuous theory we start from Eq. (3) of this Letter, which defines the jpd of eigenvalues of the Wishart ensemble, we first rewrite it in exponential form and rescale all the eigenvalues as  $\lambda_i \rightarrow \beta N \lambda_i$ . This results in  $P_\beta(\boldsymbol{\lambda}) \propto \exp(-\beta N^2 E[\{\boldsymbol{\lambda}\}])$ , with

$$E[\{\boldsymbol{\lambda}\}] = \frac{1}{2N} \sum_{j=1}^N \lambda_j - \frac{\alpha}{2N} \sum_{j=1}^N \ln \lambda_j - \frac{1}{2N^2} \sum_{j \neq k} \ln |\lambda_j - \lambda_k|,$$

where we have kept only the relevant terms. Next, we rename  $\lambda_1$  as  $\xi$  and separate it from the rest of eigenvalues which, abusing notation, are denoted again as  $\boldsymbol{\lambda} = (\lambda_2, \dots, \lambda_N)$ , viz.

$$E[\{\xi, \boldsymbol{\lambda}\}] = \frac{1}{2N} \left( \xi + \sum_{j=2}^N \lambda_j \right) - \frac{\alpha}{2N} \left( \ln \xi + \sum_{j=2}^N \ln \lambda_j \right) - \frac{1}{2N^2} \left( \sum_{j \neq k} \ln |\lambda_j - \lambda_k| + 2 \sum_{j=2}^N \ln |\xi - \lambda_j| \right)$$

This expression for  $E[\{\xi, \lambda\}]$  can now be expressed as a continuous theory in the following form:

$$S[\rho(\lambda; \lambda)] \equiv E[\{\xi, \rho(\lambda; \lambda)\}] = \frac{1}{2} \int d\lambda \rho(\lambda; \lambda) - \frac{\alpha}{2} \int d\lambda \rho(\lambda; \lambda) \ln \lambda - \frac{1}{2} \iint d\lambda d\lambda' \rho(\lambda; \lambda) \rho(\lambda'; \lambda) \ln |\lambda - \lambda'| + \mathcal{O}(N^{-1}) \quad (0.1)$$

where we have introduced the normalized one-point function

$$\rho(\lambda; \lambda) = \frac{1}{N-1} \sum_{i=2}^N \delta(\lambda - \lambda_i).$$

It is important to note that the dependence on  $\xi$  does not appear in the leading terms of Eq. (0.1) and therefore will be irrelevant for large  $N$ . That is why we write the energy as  $S[\rho(\lambda; \lambda)]$ , that is, only dependent on the density  $\rho(\lambda)$ . Going back to our expression for the cumulative function we write

$$\mathcal{P}[\kappa < x] \propto \int \mathcal{D}[\rho] \int_0^\infty d\xi \left[ \int_{\lambda_1}^{\kappa \lambda_1} d\lambda_2 \cdots \int_{\lambda_1}^{\kappa \lambda_1} d\lambda_N \right] e^{-\beta N^2 S[\rho(\lambda)]} \delta_{(F)} \left( \rho(\lambda) - \frac{1}{N-1} \sum_{i=2}^N \delta(\lambda - \lambda_i) \right)$$

Finally, using a Fourier representation for the functional Dirac delta, integrating its corresponding Lagrange multiplier, and ignoring subleading terms, we end up with the following expression:

$$\mathcal{P}[\kappa < x] \propto \int_0^\infty d\xi \int \mathcal{D}[\rho, C] e^{-\beta N^2 S[\rho, C; \xi, x\xi]}$$

with

$$S[\rho, C; L, U] = \frac{1}{2} \int_L^U d\lambda \lambda \rho(\lambda) - \frac{\alpha}{2} \int_L^U d\lambda \rho(\lambda) \ln \lambda - \frac{1}{2} \iint_L^U d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \ln |\lambda - \lambda'| - C \left( \int_L^U d\lambda \rho(\lambda) - 1 \right) + \mathcal{O}(N^{-1}). \quad (0.2)$$

as in Eq. (7) of the Letter.

### Solution of integral equation Eq. (8)

The integral equation Eq. (8) of the Letter can be solved resorting to the theory of so-called Carleman equation, which is an integral equation of type

$$\int_a^b \ln |x - t| y(t) dt = f(x). \quad (0.3)$$

In our case we have  $a = L$ ,  $b = U$ ,  $y(t) = \rho^*(y)$ , and  $f(x) = \frac{x}{2} - \frac{\alpha}{2} \ln x - C$ . For our purposes, the general solution of (0.3) reads

$$y(x) = \frac{1}{\pi^2 \sqrt{(x-a)(b-x)}} \left[ \int_a^b \frac{\sqrt{(t-a)(b-t)} f'(t) dt}{t-x} + B \right].$$

Recalling the following integral results:

$$\int_a^b \frac{\sqrt{(t-a)(b-t)}}{t-x} dt = \pi \left( \frac{b+a}{2} - x \right), \quad \int_a^b \frac{\sqrt{(t-a)(b-t)}}{t(t-x)} dt = -\frac{\pi}{x} \left( x - \sqrt{ab} \right),$$

we obtain the following expression for  $\rho^*(x)$ :

$$\rho^*(x) = \frac{1}{2\pi \sqrt{(x-a)(b-x)}} \left[ \frac{b+a}{2} - x + \frac{\alpha}{x} \left( x - \sqrt{ab} \right) + B \right],$$

where  $B$  is determined by looking for the physical solution. Imposing normalization and using the following results

$$\begin{aligned} \int_a^b dx \frac{1}{2\pi\sqrt{(x-a)(b-x)}} &= \frac{1}{2}, & \int_a^b dx \frac{1}{2\pi x\sqrt{(x-a)(b-x)}} &= \frac{1}{2\sqrt{ab}}, \\ \int_a^b dx \frac{x}{2\pi\sqrt{(x-a)(b-x)}} &= \frac{a+b}{4}, \end{aligned}$$

we obtain that  $B = 2$ . Finally, setting  $a = L$  and  $b = U$  we write

$$\rho^*(\lambda) = \frac{(x_+(L, U) - \lambda)(\lambda - x_-(L, U))}{2\pi\lambda\sqrt{(U - \lambda)(\lambda - L)}} \mathbb{1}_{[L, U]}(\lambda), \quad (0.4)$$

where  $x_{\pm}(L, U)$  are the roots of the polynomial

$$\lambda^2 - \lambda \left( \frac{L+U}{2} + \alpha + 2 \right) + \alpha\sqrt{LU} = 0$$

Writing (0.4) in this way is most convenient as we can see that the physical normalized solution is the one such that  $x_-(L, U) \leq L < U \leq x_+(L, U)$ . This implies that the Coulomb fluid has four phases:

- $x_+(L, U) = U$  and  $x_-(L, U) = U$ . In this case we notice that the spectral density becomes the MP law (Region II in Fig. 2):

$$\rho^*(\lambda) = \frac{\sqrt{(U - \lambda)(\lambda - L)}}{2\pi\lambda}, \quad L = z_-, \quad U = z_+.$$

- $x_+(L, U) = U$  and  $x_-(L, U) < L$ . In this case the spectral density becomes (Region III of Fig. 2):

$$\rho^*(\lambda) = \frac{1}{2\pi} \sqrt{\frac{U - \lambda}{\lambda - L}} \left[ \frac{\lambda - \alpha\sqrt{L/U}}{\lambda} \right], \quad L(U) = \frac{1}{U} \left[ \alpha + \sqrt{(\alpha - U)^2 - 4U} \right]^2, \quad U \geq \left( 1 + \frac{1}{\sqrt{c}} \right)^2.$$

Note that in this case we take as free parameter the position of the upper barrier  $U$ , instead of the lower barrier, as the latter will result into a third degree polynomial to solve.

- $x_-(L, U) = L$  and  $x_+(L, U) > U$ . In this case the spectral density becomes (Region IV of Fig. 2):

$$\rho^*(\lambda) = \frac{1}{2\pi} \sqrt{\frac{\lambda - L}{U - \lambda}} \left[ \frac{\alpha\sqrt{U/L} - \lambda}{\lambda} \right], \quad U(L) = \frac{1}{L} \left[ \alpha - \sqrt{(\alpha - L)^2 - 4L} \right]^2, \quad L \leq \left( 1 - \frac{1}{\sqrt{c}} \right)^2,$$

where in this case we take the lower barrier  $L$  as our free parameter.

- In this case we have  $x_+(L, U) > U$  and  $x_-(L, U) < L$  and (Region I of Fig. 2):

$$\rho^*(\lambda) = \frac{[x_+(L, U) - \lambda][\lambda - x_-(L, U)]}{2\pi\lambda\sqrt{(U - \lambda)(\lambda - L)}}$$

with  $x_{\pm}(L, U)$  are the solutions

$$x_{\pm}(L, U) = \frac{1}{2} \left[ \left( \frac{L+U}{2} + \alpha + 2 \right) \pm \sqrt{\left( \frac{L+U}{2} + \alpha + 2 \right)^2 - 4\alpha\sqrt{LU}} \right]$$

#### Left rate function

To derive the left rate function we need to evaluate the action at the solution of the saddle point, that is  $S[\rho^*, C; L, U]$ . It is possible to get rid of the double integral appearing (0.2) by taking the saddle-point equation, multiplying by  $\rho^*(\lambda)$ , and



integrating  $\lambda$  over the interval  $[L, U]$ . The constant  $C$  is then determined taking the value of  $\lambda = U$  at the saddle-point equation. All in all, the action at the saddle point takes the following form

$$S[\rho^*, C; L, U] = \frac{1}{4} \int_L^U d\lambda \rho^*(\lambda) (\lambda - \alpha \ln \lambda) + \frac{1}{2} C(L, U)$$

with

$$C(L, U) = \frac{U}{2} - \frac{\alpha}{2} \ln U - \int_L^U d\lambda \rho^*(\lambda) \ln |U - \lambda|$$

Now we report the result for the following integrals:

$$\begin{aligned} \int_L^U d\lambda \rho^*(\lambda) \lambda &= \frac{1}{4} (L + U) \left( \alpha + \frac{L + U}{2} + 2 \right) - \frac{1}{2} \alpha \sqrt{LU} - \frac{1}{16} (3L^2 + 2LU + 3U^2) \\ \int_L^U d\lambda \rho^*(\lambda) \ln \lambda &= \frac{1}{4} \left( 8(\alpha + 1) \ln(\sqrt{L} + \sqrt{U}) - 2\alpha \ln(L) - 2\alpha \ln(U) - 8(\alpha + 1) \ln(2) + 2\sqrt{LU} - L - U \right) \\ \int_L^U d\lambda \rho^*(\lambda) \ln |U - \lambda| &= \frac{1}{4} \left( -2\alpha \ln \left( \frac{U(-2\sqrt{LU} + L + U)}{U - L} \right) + 2(\alpha + 2) \ln(U - L) - 4(\alpha + 2) \ln(2) - L + U \right) \end{aligned}$$

Gathering all these results and after simplifying we obtain the final expression for the action

$$\begin{aligned} S[\rho^*, C; L, U] &= -\frac{1}{64} \left( -2L(4\alpha + U + 8) + 16\alpha\sqrt{LU} - 8\alpha \left( -4(\alpha + 2) \ln(\sqrt{L} + \sqrt{U}) + \alpha \ln(L) + \alpha \ln(U) \right) \right. \\ &\quad \left. - 8(\alpha + 2)U - 32(\alpha(\alpha + 2) + 2) \ln(2) + L^2 + 32 \ln(U - L) + U^2 \right) \end{aligned}$$

Two more steps need to be done to get the left rate function: the first one is to take  $L = \xi$  and  $U = x\xi$  and minimize  $S[\rho^*, C; \xi, x\xi]$  with respect to  $\xi$ . This gives

$$\frac{d}{d\xi} S[\rho^*, C; \xi, x\xi] \Big|_{\xi=\xi^*} = 0 \Rightarrow \xi^*(x) = 4 \frac{1 + \alpha}{(1 + \sqrt{x})^2}.$$

The second step is to obtain the normalization factor by taking the limit  $x \rightarrow \infty$  in the action. This, in the Coulomb fluid picture, corresponds to  $x \rightarrow \langle \kappa \rangle$ . By noting that  $\xi^*(\langle \kappa \rangle) = z_-$  and that  $\rho^* \rightarrow \rho_{\text{mp}}$ , the left rate function  $\Phi_-(x) = S[\rho^*, C; \xi^*, x\xi^*] - S[\rho_{\text{mp}}, C; z_-, z_+]$  gives

$$\begin{aligned} \Phi_-(x) &= \frac{1}{8(\sqrt{x} + 1)^2} \left[ 4\alpha^2 (x \ln(2) + \sqrt{x} \ln(4) - \ln(\sqrt{x} + 1)) + \alpha^2 (\sqrt{x} + 1)^2 \ln(x) \right. \\ &\quad - 2(x + 2\sqrt{x}) (2\alpha^2 \ln(\sqrt{x} + 1) + \ln(\alpha + 1)) \\ &\quad + 4\alpha^2 (\sqrt{x} + 1)^2 \operatorname{arccoth}(2\alpha + 1) + \alpha^2 \ln(16) + 2(\alpha + 1) (\alpha x - 2(\alpha + 2)\sqrt{x} + \alpha) \\ &\quad \left. - 2 \ln(\alpha + 1) + 8(\sqrt{x} + 1)^2 \operatorname{arccoth}(\sqrt{x}) \right]. \end{aligned}$$

Finally, after some algebra,  $\Phi_-(x)$  can be rewritten as reported in Eq. (10), and in terms of the functions  $f_1^{(\alpha)}(\omega)$  and  $f_2^{(\alpha)}(\omega)$

$$f_1^{(\alpha)}(\omega) = \frac{2(\alpha + 1)(\alpha(\omega - 2)^2 - 4\omega + 4)}{\omega^2}, \quad f_2^{(\alpha)}(\omega) = \left( \frac{4(\alpha + 1)(\omega - 1)}{\alpha\omega^2} \right)^{2\alpha^2} \left( \frac{\omega}{\sqrt{1 + \alpha}(\omega - 2)} \right)^4.$$

*Square case: Edelman's result*

In this section, we show how one recovers Edelman's results for the condition number in the *square* matrix case ( $N = M$ ). Two of the results contained in Edelman's paper [19] are as follows. For  $\beta = 1$ ,

$$\lim_{N \rightarrow \infty} \mathcal{P}[\hat{\kappa}/N < x] = e^{-2/x - 2/x^2}, \quad (0.5)$$

and for  $\beta = 2$

$$\lim_{N \rightarrow \infty} \mathcal{P}[\hat{\kappa}/N < x] = e^{-4/x^2}, \quad (0.6)$$

where  $\hat{\kappa} = \sqrt{\kappa}$  (as defined in our Letter). Thus,  $\mathcal{P}[\hat{\kappa} < Nx] = \mathcal{P}[\kappa < N^2 x^2]$ . Our result reads  $\mathcal{P}[\kappa < x] \approx \exp(-\beta N^2 \Phi_-(x))$ . So all we have to do is to compute  $\Phi_-(N^2 x^2)$  for  $\alpha = 0$  to leading order in  $N$ . We get

$$N^2 \Phi_-(N^2 x^2) \sim \frac{2}{x^2}, \quad \text{for } N \rightarrow \infty, \quad (0.7)$$

implying

$$\mathcal{P}[\hat{\kappa} < Nx] \approx \exp\left(-\frac{2\beta}{x^2}\right), \quad (0.8)$$

which correctly reproduce Edelman's result for  $\beta = 2$  and his leading term in  $N$  for  $\beta = 1$ . The missing contribution in the latter case corresponds to a  $\mathcal{O}(N)$  correction in the Coulomb gas approach not captured by our theory.

### Right rate function

First, we give the definition of the functions  $g^{(\alpha)}(x)$  and  $h^{(\alpha)}(x)$  appearing in Eq. (13) of the Letter, as

$$g^{(\alpha)}(x) = \Xi\left(\sqrt{\varpi(x, t) + 1}, t\right), \quad (0.9)$$

$$h^{(\alpha)}(x) = \frac{\sqrt{\varpi(x, t)}}{1 + \sqrt{\varpi(x, t) + 1}}, \quad (0.10)$$

$$\varpi(y, \tau) = \frac{4\tau(1+y)}{(\tau-1)^2(y - \langle \kappa \rangle)}, \quad (0.11)$$

$$\Xi(r, s) = \frac{s^2(r+1)^2 - (r-1)^2}{s^2(r-1)^2 - (r+1)^2}, \quad (0.12)$$

where  $t = \sqrt{1 + \alpha}$ .

Next, we derive the difference in Coulomb fluid free energy when the minimum eigenvalue is pinned below the left edge  $z_-$  of the MP law and the maximum eigenvalue above the right edge  $z_+$  (at the positions  $L$  and  $U$  respectively). We want to estimate  $\Delta E = E(L, \lambda_2, \dots, \lambda_{N-1}, U) - E_{\text{mp}}(\lambda_1, \dots, \lambda_N)$ , where the initial energy  $E_{\text{mp}}$  is the one corresponding to the unperturbed MP configuration. Recall that the Coulomb energy is

$$E(\lambda_1, \dots, \lambda_N) = \frac{1}{2} \sum_{i=1}^N \lambda_i - \left(\frac{\beta}{2}(\alpha N + 1) - 1\right) \sum_{i=1}^N \ln \lambda_i - \frac{\beta}{2} \sum_{i \neq j} \ln |\lambda_i - \lambda_j|. \quad (0.13)$$

Then we have (after pinning the extreme particles)

$$\begin{aligned} E(L, \lambda_2, \dots, \lambda_{N-1}, U) &= \frac{L+U}{2} + \frac{1}{2} \sum_{i=2}^{N-1} \lambda_i - \left(\frac{\beta}{2}(\alpha N + 1) - 1\right) (\ln L + \ln U) - \left(\frac{\beta}{2}(\alpha N + 1) - 1\right) \sum_{i=2}^{N-1} \ln \lambda_i \\ &\quad - \beta \sum_{2 \leq i < j \leq N-1} \ln |\lambda_i - \lambda_j| - \beta \sum_{i=2}^{N-1} \ln |L - \lambda_i| + \beta \sum_{i=2}^{N-1} \ln |\lambda_i - U| - \beta \ln |U - L|. \end{aligned} \quad (0.14)$$

Rescaling the eigenvalues  $\lambda_i \rightarrow \beta N \lambda_i$ , and recalling that  $\lambda_{\min} \sim z_-$ ,  $\lambda_{\max} \sim z_+$  we write the energy difference as  $\Delta E(L, U) = \beta N \Delta e(L, U)$  with

$$\Delta e(L, U) = \frac{(L - z_-) + (U - z_+)}{2} - \frac{\alpha}{2} \ln \frac{LU}{z_- z_+} + \int_{z_-}^{z_+} d\eta \rho_{\text{mp}}(\eta) \ln \left| \frac{(L - \eta)(\eta - U)}{(z_- - \eta)(\eta - z_+)} \right|, \quad (0.15)$$

By replacing  $L = \xi$  and  $U = x\xi$ , and denoting the corresponding energy difference as  $\Delta e(\xi, \xi x)$ , we obtain Eq. (11) of the Letter.

The integrals appearing in this expression can be evaluated analytically. We obtain the following results

$$\begin{aligned} \int_{z_-}^{z_+} d\eta \rho_{\text{mp}}(\eta) \ln \left| \frac{\xi - \eta}{z_- - \eta} \right| &= \frac{1}{2} \left( 2(2 + \alpha) \operatorname{arcsinh} \left( \frac{1}{2} \sqrt{\frac{z_- - \xi}{\sqrt{1 + \alpha}}} \right) + \alpha \left( \ln \left( \frac{z_-}{\xi} \right) \right. \right. \\ &\quad \left. \left. - 2 \operatorname{arctanh} \left( \sqrt{\frac{z_+(z_- - \xi)}{z_-(z_+ - \xi)}} \right) \right) - z_- + \xi + \sqrt{(z_- - \xi)(z_+ - \xi)} \right) \\ \int_{z_-}^{z_+} d\eta \rho_{\text{mp}}(\eta) \ln \left| \frac{\eta - \xi x}{\eta - z_+} \right| &= \frac{1}{2} \left( \alpha \left( \ln \frac{z_+}{\xi x} - 2 \operatorname{arctanh} \left( \sqrt{\frac{z_-(\xi x - z_+)}{z_+(\xi x - z_-)}} \right) \right) \right. \\ &\quad \left. + 2(2 + \alpha) \operatorname{arcsinh} \left( \frac{1}{2} \sqrt{\frac{\xi x - z_+}{\sqrt{1 + \alpha}}} \right) + \xi x - z_+ - \sqrt{(\xi x - z_-)(\xi x - z_+)} \right). \end{aligned} \quad (0.16)$$

After some algebra one can arrive to the intermediate expression

$$\begin{aligned} \Delta e(\xi, \xi x) &= -\frac{1}{2} \left[ -2\alpha \left( \operatorname{arctanh} \left( \sqrt{\frac{z_+(z_- - \xi)}{z_-(z_+ - \xi)}} \right) + \operatorname{arctanh} \left( \sqrt{\frac{z_-(\xi x - z_+)}{z_+(\xi x - z_-)}} \right) \right) \right. \\ &\quad \left. + 2(2 + \alpha) \left( \operatorname{arcsinh} \left( \frac{1}{2} \sqrt{\frac{x\xi - z_+}{\sqrt{1 + \alpha}}} \right) + \operatorname{arcsinh} \left( \frac{1}{2} \sqrt{\frac{z_- - \xi}{\sqrt{1 + \alpha}}} \right) \right) \right. \\ &\quad \left. - \sqrt{(x\xi - z_-)(x\xi - z_+)} + \sqrt{(z_- - \xi)(z_+ - \xi)} \right]. \end{aligned} \quad (0.17)$$

Recall that the tail cumulative will be given by  $\mathcal{P}[\kappa > x] \approx \exp[-\beta N \Phi_+(x)]$ , where  $\Phi_+(x) = \Delta e(\xi^*(x), \xi^*(x)x)$  so all is left to do is to minimize  $\Delta e$  with respect to  $\xi$ . One can convince oneself that this is given by

$$\left. \frac{d}{d\xi} \Delta e(\xi, \xi x) \right|_{\xi=\xi^*} = 0 \Rightarrow \xi^*(x) = \frac{2(2 + \alpha)}{x + 1} \quad (0.18)$$

Plugging this result back to  $\Delta e(\xi^*(x), \xi^*(x)x)$  and after some algebra we get to the result (12) of this Letter.

### Order of phase transitions

An analysis of the derivatives of the rate functions will reveal the type of transition as they approach  $\langle \kappa \rangle$ . We obtain

$$\begin{aligned} \Phi_-(x) &= -\frac{(-1 + \sqrt{1 + \alpha})^8}{96\sqrt{1 + \alpha}(1 + \sqrt{1 + \alpha})^4} (\langle \kappa \rangle - x)^3 + \dots \\ \Phi_+(x) &= \frac{\sqrt{2}^4 \sqrt{\alpha + 1} (\sqrt{\alpha + 1} - 1)^4}{3\sqrt{\alpha + 2} (\sqrt{\alpha + 1} + 1)^2} (x - \langle \kappa \rangle)^{3/2} + \dots \end{aligned} \quad (0.19)$$

### Numerical simulations for the left rate function

Let us recall the expression for the cumulative distribution

$$\mathcal{P}[\kappa < x] = \frac{1}{(N - 1)!} \int_0^\infty d\lambda_1 \left[ \int_{\lambda_1}^{x\lambda_1} d\lambda_2 \cdots \int_{\lambda_1}^{x\lambda_1} d\lambda_N P_\beta(\boldsymbol{\lambda}) \right] \quad (0.20)$$

Suppose we were to implement a standard Monte Carlo simulation of the Coulomb fluid. After relaxation has been achieved, we would simply use the Monte Carlo Markov Chain to estimate  $\mathcal{P}[\kappa < x]$  by keeping a record of the number of states such that  $\lambda_i \in [\lambda_1, x\lambda_1]$  for  $i = 2, \dots, N$  divided by the total number of recorded states, that is

$$\mathcal{P}[\kappa < x] \propto \frac{\#(\{\lambda_1, \dots, \lambda_N\} | \lambda_i \in [\lambda_1, x\lambda_1], i = 2, \dots, N)}{\#(\{\lambda_1, \dots, \lambda_N\})} \quad (0.21)$$

There is nothing wrong with this procedure except that it is extremely inefficient: for atypical values of  $x$  it will be very unlikely to find the Coulomb fluid in that configuration, and a huge number of samples will be needed to have reliable statistics. There is a smart way around this, though. One notices that it is very easy to obtain the samples  $\#(\{\lambda_1, \dots, \lambda_N\} | \lambda_i \in [\lambda_1, x\lambda_1], i = 2, \dots, N)$  by simulating the fluid directly with such constraints. The tricky part is to realize how to get the cumulative by using this subset of samples. We write

$$\mathcal{P}[\kappa < y] = \mathcal{P}[\kappa < y | y \leq x] K_x \quad (0.22)$$

where  $K_x \equiv \mathcal{P}[y \leq x]$  depends only on  $x$ . The idea is then to simulate the Coulomb fluid such that  $\lambda_i \in [\lambda_1, x\lambda_1]$  and to calculate the cumulative  $\mathcal{P}[\kappa < y | y \leq x]$ . Obviously the cumulative will be concentrate mainly in the interval  $y \in [x - \delta, x]$  for some small  $\delta$ . We can then use this to estimate the derivative of the left rate function, getting rid of the constant  $K_x$ , that is

$$\Phi'_-(y) = -\frac{1}{\beta N^2} \frac{d}{dy} \ln \mathcal{P}[\kappa < y | y \leq x] \quad (0.23)$$

Therefore, the algorithmic procedure is the following:

1. Choose a value of  $x$ .
2. Simulate the Coulomb fluid with barriers such that  $\lambda_i \in [\lambda_1, x\lambda_1]$  and let it thermalize.
3. Construct the histogram of the cumulative  $\mathcal{P}[\kappa < y | y \leq x]$  of this confined fluid.
4. Estimate the derivative  $\Phi'_-(y)$  using the constructed histogram

Then by scanning the values of  $x$  we obtain the derivative of the left rate function in the interval  $x \in [1, \langle \kappa \rangle]$ .

---